



Modeling of route planning system based on Q value-based dynamic programming with multi-agent reinforcement learning algorithms



Mortaza Zolfpour-Arokhlo^{a,*}, Ali Selamat^b, Siti Zaiton Mohd Hashim^b, Hossein Afkhami^c

^a Department of Computer Engineering, Sepidan Branch, Islamic Azad University, Sepidan, Iran

^b Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia

^c Department of Electronics and Power, Sepidan Branch, Islamic Azad University, Sepidan, Iran

ARTICLE INFO

Article history:

Received 29 April 2013

Received in revised form

1 December 2013

Accepted 5 January 2014

Available online 31 January 2014

Keywords:

Route planning system (RPS)

Multi-agent system (MAS)

Multi-agent reinforcement learning (MARL)

Q-learning

Traffic congestion

ABSTRACT

In this paper, a new model for a route planning system based on multi-agent reinforcement learning (MARL) algorithms is proposed. The combined Q-value based dynamic programming (QVDP) with Boltzmann distribution was used to solve vehicle delay's problems by studying the weights of various components in road network environments such as weather, traffic, road safety, and fuel capacity to create a priority route plan for vehicles. The important part of the study was to use a multi-agent system (MAS) with learning abilities which in order to make decisions about routing vehicles between Malaysia's cities. The evaluation was done using a number of case studies that focused on road networks in Malaysia. The results of these experiments indicated that the travel durations for the case studies predicted by existing approaches were between 0.00 and 12.33% off from the actual travel times by the proposed method. From the experiments, the results illustrate that the proposed approach is a unique contribution to the field of computational intelligence in the route planning system.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Route planning systems (RPS) are one of several types of traffic information systems that offer routes that solve traffic problems. RPS provides optimum route solutions and traffic information (Ji et al., 2012) prior to a trip in order to help drivers arrive at their destination as quickly as possible. Route planning problems can be solved by determining the shortest paths using a model of the transportation network (Kosicek et al., 2012; Geisberger, 2011). The drivers of vehicles with different solutions available to them need quick updates when there are changes in road network conditions. Unfortunately, the task of efficiently routing vehicles in the route planning (Suzuki et al., 1995; Pellazar, 1998; Stephan and Yunhui, 2008) has not been emphasized enough in recent studies. Multi-agent systems (MAS) are a group of autonomous, interacting entities sharing a common environment, which they perceive with sensors and upon which they act with actuators (Shoham and Leyton-Brown, 2008; Vlassis, 2007). MASs are applied in a variety of areas, including robotics, distributed control, resource management, collaborative decision support systems (DSS), and data mining (Bakker et al., 2005; Riedmiller et al., 2000). They can be used as a natural way of operating on a system, or they may provide an alternative perspective for centralized systems. For

instance, in robotic teams, controlling authority is naturally distributed between the robots. Reinforcement learning (RL), which allows (Tesauro et al., 2006; Lucian et al., 2010; Bakker and Kester, 2006) learning provides an environment for learning how to plan, what to do and how to map situations (Jie and Meng-yin, 2003) to actions, and how to maximize a numerical reward signal. In an RL, the learner is not told which actions to take, as is common in most forms of machine learning. Instead, the learner must discover through trial and error, which actions yield the most rewards. In the most interesting and challenging cases, actions affect not only the immediate rewards but also the next station or subsequent rewards. The characteristics of trial and error searches and delayed reward are two important distinguishing features of RL, which are defined not by characterizing learning methods, but by characterizing a learning problem. Any method that is suitable for problem solving is considered to be an RL method. An agent must be able to sense the state of the environment, and be able to take actions that affect the environment. The agent must also have goals related to the state of the environment. In other words, an RL agent learns by interacting with its dynamic environment. The agent perceives the state of the environment and takes actions that cause the environment to transit into a new state. A scalar reward signal evaluates the quality of each transition, and the agent must maximize the cumulative rewards along the course of interaction. RL system feedback reveals if an activity was beneficial and if it meets the objectives of a learning system by maximizing expected rewards over a period of time (Shoham and Leyton-Brown, 2008;

* Corresponding author.

E-mail address: zolfpour@gmail.com (M. Zolfpour-Arokhlo).

Busoniu et al., 2005). The reward (RL feedback) is less informative than it is in supervised learning, where the agent is given the correct actions to perform (Busoniu et al., 2010). Unfortunately, information regarding correct actions is not always available. RL feedback, however, is more informative than un-supervised learning feedback, where no explicit comments are made regarding performance. Well-understood, provably convergent algorithms are available for solving single-agent RL tasks. MARL faces challenges that are different from the challenges faced by single-agent RL such as convergence, high dimensionality, and multiple equilibria. In route planning processes (Schultes, 2008; Wedde et al., 2007), route planning should take into account traveler responses and even use these responses as its guiding strategy, and realize that a traveler's route choice behavior will be affected by the guidance control decision-making of route planning. On the other hand, the results of a traveler's choice of routes will determine network conditions and react to the guidance control decision-making of route planning (Dong et al., 2007; Yu et al., 2006). Reduced vehicle delays could be achieved by examining several conditions that affect transportation network studying the weight of transport environmental conditions (Tu, 2008; Zegeye, 2010). These conditions include: weather, traffic information, road safety (Camiel, 2011), accidents, seasonal cyclical effects (such as time-of-day, day-of-the-week and month) and cultural factors, population characteristics, traffic management (Tampere et al., 2008; Isabel et al., 2009; Almejalli et al., 2009; Balaji et al., 2007; Chang-Qing and Zhao-Sheng, 2007) and traffic mix. Regarding these variables can be used to provide a priority trip plan to vehicles for drivers. Increasingly, information agent-based route planning (Gehrke and Wojtusiak, 2008) and transportation system (Chowdhury and Sadek, 2003) applications are being developed. The goal of this study is to enable multiple agents to learn suitable behaviors in a dynamic environment using an RL that could create cooperative (Lauer and Riedmiller, 2000; Wilson et al., 2010) behaviors between the agents that had no prior-knowledge. The physical accessibility of traffic networks provided selective negotiation and communication between agents and simplified the transmission of information. MAS was suited to solve problems in complex systems because it could quickly generate routes to meet the real-time requirements of those systems (Shi et al., 2008). Our survey review indicated that MAS methods and techniques were applied in RPSs (Flinsenberg, 2004; Dellinger, 2009) including dynamic routing, modeling, simulation, congestion control and management, traffic control, decision support, communication, and collaboration. The main challenge faced by RPS was directing vehicles to their destination in a dynamic real-time RTN situation while reducing travel time and enabling the efficient use of the available capacity on the RTN (Khanjary and Hashemi, 2012). To avoid congestion and to facilitate travel, drivers need traffic direction services. These services lead to more efficient traffic flows on all transport networks, resulting in reduced pollution and lower fuel consumption. Generally, these problems are solved using a fast path planning method and it seems to be an effective approach to improve RPS. For example, RPS is used in the following areas: traffic control, traffic engineering, air traffic control (Volf et al., 2011), trip planning, RTN, traffic congestion, traffic light control, traffic simulation, traffic management, urban traffic, traffic information (Ji et al., 2012; Khanjary and Hashemi, 2012) and traffic coordination (Arel et al., 2010). The main thrust of using RPS in these situations is to compare the shortest path algorithms with Dijkstra's algorithm. Currently, there are various successful algorithms for shortest path problems that can be used to find the optimal and the shortest path in the RPS. The research will contribute to

- Using the ability of learning models to propose several route alternatives.
- Predicting the road and environmental conditions for vehicle trip planning.
- Providing high quality travel planning service to the vehicle drivers.

This paper consists of seven sections. Section 2 describes the meaning of RPS based on MARL and related works. This will be followed by a definition of the RPS based on the MARL problem (Section 3). Section 4 will present the MARL proposed for RPS. Section 5 discusses the experimental method used in this study. Section 6 presents the results, comparisons, and evaluations. Finally, the paper is concluded in the last section.

2. Related works

RPS based on multi-agent decisions is described as a tool in transportation planning (Dominik et al., 2011). The agent chooses among, competing vendors, distributes orders to customers, manages inventory and production, and determines price based on a perfect competitive behavior. The concept of an agent-based trip planning in transportation was studied by Yunhui and Stephan (2007) to reduce the negative effects of disruptive events. The RTN environment is highly dynamic, complex and has many constraints. In this study, a major effort was made to improve MAS solutions and algorithms to use in a simulated transportation environment (Kovalchuk, 2009). Consequently, each agent managed a specific function of the transportation environment network and shared information about other agents. Communication between agents was based on local information. As discussed above, the agents had their own problem solving capabilities and were able to interact with each other in order to reach a goal. Interaction could occur between agents (agent-agent interaction) and between agents and their environment. In the environmental sciences, MAS is often used as a platform (Jones, 2010; Robu et al., 2011) for space-time dynamics. The characteristics of the agents encountered also differed, ranging from simple reactive agents to more intelligent agents that show a limited capacity for reasoning and making decisions as noted by Ligtenberg (2006). RPS integrates technologies such as information technology, networks, communications, and computer science (Shimohara et al., 2005). A new approach in RPS is being studied that can save space and costs for cars at intersections or junctions (Vasirani and Ossowski, 2009). Chen and Cheng (2010) surveyed different applications of agent technology in traffic modeling and simulations. They emphasized the importance of agent technology for improvement of the RTN and for traffic system performance. Ruimin and Qixin (2003) conducted an extensive study of coordination controls and traffic flows that focused on agent-based technology. All of these studies used models based on agent flexibility to improve traffic management. The aim of this study was to propose a novel RPS approach based on MAS and MARL that could perform better than previous RTN methods such as those discussed (Adler et al., 2005; Wu et al., 2011a), in terms of accuracy and real-time performance. The results obtained from this approach were compared with Dijkstra's algorithm, which was used to find the best and optimal path between the origin and destination nodes in a directed transportation graph. Agents were defined as the specific functional performance of a set of inputs in Uppin (2010). The key problem faced by an agent was that of deciding which action it should perform in order to best satisfy its design objectives. Using suitable agent architectures facilitated agent decision making. Agent architecture is software architecture for decision-making systems that are embedded in an environment (Michael, 2009).

- Modeling a new route planning system based on QVDP with the MARL algorithm.

Download English Version:

<https://daneshyari.com/en/article/380661>

Download Persian Version:

<https://daneshyari.com/article/380661>

[Daneshyari.com](https://daneshyari.com)